STAC HOT SPOT AREAS:
A STATISTICAL TOOL FOR LAW ENFORCEMENT DECISIONS[1]

CAROLYN REBECCA BLOCK
Statistical Analysis Center,
Illinois Criminal Justice Information Authority

The last ten or fifteen years have seen a quiet revolution in criminology and criminal justice. There has been a vast improvement in the quality and quantity of data, and in the availability of that data to decision-makers. Sufficient information is now available so that basic indicators related to criminal justice issues can be measured with a degree of precision that was not only unknown a few years ago, but was even unanticipated. Contrary to conventional wisdom, we are not plagued by a dearth of information in criminal justice (Block, 1989). The problem is just the opposite. There are often so many pieces of information that it is impossible for the human mind to assimilate them, to sort them out, and to use them for tactical or crime analysis decisions before the window of opportunity has passed.

Even more recently, another revolution has not only precipitated a surge in the *amount* of information available, but also *enlarged the nature* of that information to include two-dimensional space. This technological revolution, computer mapping, has generated a need for analytical methods and techniques to make spatial information a foundation for answering practical and policy questions. One of the most common spatial analysis questions asked in law enforcement is: where are the densest concentrations of incidents or events on the map? The Hot Spot Area capability of the STAC (Spatial and Temporal Analysis of Crime) software program was developed to answer that question.

This presentation is an overview of the STAC approach to finding and describing Hot Spot Areas. It begins with a review of cluster analysis techniques as they have been applied to law enforcement. It then describes Hot Spot Area analysis in STAC, and provides an example of a Chicago violence-reduction project that uses Hot Spot Area analysis to identify street gang violence crisis areas.

---

[1]An earlier version of this paper was presented at the May, 1993, meetings on Environmental Criminology, Miami Florida, and published in the Proceedings of those meetings.

COMPUTER MAPPING: A TECHNOLOGICAL REVOLUTION IN LAW ENFORCEMENT

Computer mapping technology, coupled with the technology necessary to store and organize vast amounts of geocoded data, has expanded extremely rapidly in recent years, and there is no end to the expansion in sight. Only a few years ago, the only mapped information available to most police departments was limited to cardboard pin maps and colored plastic pins. Digitizing was (and still is) so expensive that the development of their own computerized street map was beyond the reach of most departments.[2] Cities that did have mapping capability usually housed it centrally, outside the police department (in City Planning, for example), because mapping software and hardware was complex, requiring experts to use it, and too expensive for scattered-site applications. In these cities, police access to mapping was indirect, often cumbersome and time consuming. In addition, maps developed for other city uses did not always meet law enforcement needs. (For example, we do not send a squad car to a Census tract to answer a call, but to an address.)

With the advent of accessible mapping software, PCs that can handle it, and the Census TIGER files, all this has changed. As presentations at this workshop demonstrate, the technological revolution in computer mapping means that spatial crime analysis is now a possibility, even for small police departments and neighborhood-level police districts. Mapping entrepreneurs, people with the vision to see the potential worth of automated mapping and the spirit and initiative to overcome any obstacle that might be in the way of realizing that potential, have sprung up in departments across the continent and around the world. Many of those people are with us here today.

By its very success, however, this technological revolution has generated a "data overload" of unprecedented proportions. Mapped databases can quickly grow to contain much more information than a police analyst can possibly assimilate and use for timely decisions. Computer mapping technology, no matter how sophisticated, is not enough by itself to organize the vast amounts of spatial and other information generated by the reality of daily interaction in city neighborhoods, and to sumarize that information so that it can support tactical deployment, investigation, early identification of crisis situations, and development of successful intervention strategies.

As Roncek and Maier (1991) argue and Maltz, Gordon and Friedman (1991) found in their pilot study of computer mapping in police departments, the successful analysis of spatial patterns of crime requires that empirical mapping tools be guided by theory that can link place to crime, can unravel the spatial characteristics of different types of crime, and can provide explanations for the high vulnerability of some neighborhoods or demographic groups. In other words, we must link computer mapping technology to an information-organizing framework that encompasses both law enforcement and community information. The appropriate tools for a job such as this are statistics. Statistical tools (figure 1) can summarize enormous amounts of information and organize that information to answer specific, practical questions. They can provide a way to link mapping technology and information-organizing hypotheses.

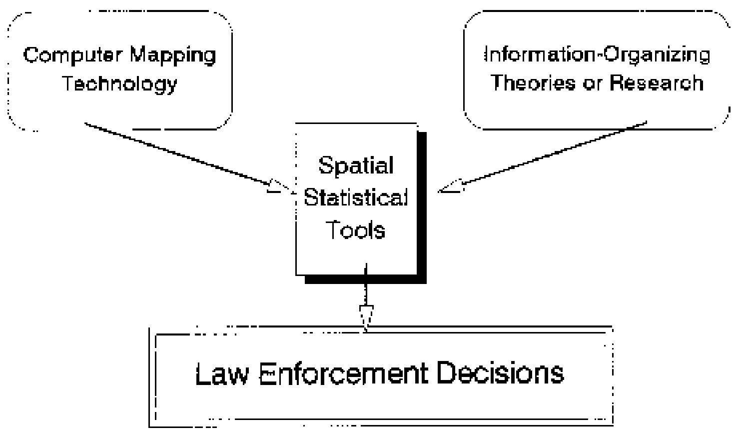
We need statistical tools to control and manage data overload, so that law enforcement and community information can become a foundation for tactical, crime analysis and policy decisions. However, as the next section will show, statistical methods for using and interpreting mapped data, especially methods applicable to practical criminal justice situations, are still in their infancy. The development of statistics for geographic analysis has been outpaced by the creation of geocoded datasets and the software to map them. Methods for identifying "Hot Spot Areas" (the densest clusters of incidents on a map) have been particularly problematic.

---

[2] Digitizing is the development of a map, such as a city street map, in which features (streets, rivers, expressways, and so on) have x-coordinates and y-coordinates for automated mapping.

Figure 1

Statistics: the Vital Link in an Information Foundation for Law Enforcement Decisions



ATTEMPTS TO IDENTIFY HOT SPOT AREAS

Maps, even a single map of one type of event in one time period, contain a huge amount of information, data that may be interpreted very differently by different observers with differing perceptions of spatial reality (cognitive maps).[3] The answer to a question such as whether or not the events within a certain area on the map are densely clustered, or just one permutation of random distribution, depends on the eye of the observer. In such situations, when the amount of information is overwhelming, a quick, efficient and objective summary of reality -- in other words, a statistical analysis -- can provide a useful guide to interpretation. Unfortunately, however, spatial analysis statistics have lagged behind the revolutionary expansion of hardware and software technology.

[3]See Smith and Patterson (1980) for a discussion of cognitive maps.

What statistics are available for finding and defining a Hot Spot Area -- the densest area on a map, an area that reflects the pattern of actual events, even if these natural clusters cross a boundary (police district or Census tract boundaries, for example) or extend along a boundary (a street, for example)? Two approaches have been commonly used to find dense clusters, analysis of density within pre-determined boundaries (areal analysis), and analysis of relative frequency of occurrence at specific places or addresses (pin map frequency distributions).[4] Either may be useful for answering certain other questions, but both suffer from serious limitations as methods for identifying and mapping a Hot Spot Area.

Areal Analysis

Statistical methods for the analysis of crime density within arbitrary areal units, such as police districts, political wards or Census tracts, have been available in automated systems for some years, but they all suffer from serious problems in interpretation, such as aggregation bias. (See Brantingham & Brantingham, 1984:15-18 for a review of these problems; examples of misinterpretation can be found in Pyle, 1974; 1976 and Kohfeld & Sprague, 1990.) Analysis within arbitrary boundaries cannot deal with a reality in which actual dense areas may cross boundary lines, occur along a boundary line, or occur in only a small but concentrated area within the arbitrary boundary. Aggregations to arbitrary areas are thus subject to what may be called an "area-unit fallacy," which is analogous to the ecological fallacy. In the ecological fallacy, area-level characteristics are mistakenly applied to individuals. In the area-unit fallacy, the aggregate characteristics of an overall area are mistakenly applied to each section or neighborhood within that area.

Even though aggregate data within arbitrary boundaries are not an appropriate basis for an analysis of event clusters, sometimes the only data available may be area-level. To handle this situation, techniques have been developed to approximate clusters of points with aggregate area information. In *small area* analysis, the areas analyzed are so small relative to the scale of the map that they begin to lose their two-dimensional aspect, and seem to the observer to become points instead of areas. For example, Roncek and Montgomery (elsewhere in this volume) analyze block-level data across cities. However, this begs the question; even if the small areas are seen as points, there is still no systematic way to identify clusters of these "points."

Crime gradient mapping (see Joelson & Fishbine, 1980: 251-255, and Rengert or LeBeau, elsewhere in this volume) is another technique to deal with area-unit data. Crime gradients are isopleths calculated by connecting the centroids of areas with similar rates.[5] Currently-available software such as SYSTAT and Idrisi make it easy to produce topographical (three dimensional) maps of crime gradients. However, because they are based on area data, these maps suffer from the limitations just discussed. They do not really represent clusters of points, but rather the characteristics of the *aggregated* points within each area boundary.

Both the isopleth maps developed by Curtis (1974:119-158) and studies of "population potential" (Choldin & Roncek, 1976; Felson, 1986) use area rates to define central points (centroids) on the map, which are then linked. However, because the centroids are based on rates calculated within arbitrary spatial boundaries, not on specific address or pin distributions, such methods are subject to the same problems in misinterpretation

---

[4]In the nomenclature of cartography, maps based on areal units are referred to as "choropleth" maps. This presentation substitutes the less technical term "areal" where possible. It also uses "pin map" or "point" as synonyms for address-based information, because many of the events that must be mapped in a law enforcement application do not occur at addresses. For example, the location of the body of a homicide victim may be in a river, underneath a viaduct, along railroad tracks, or in the middle of a large park or parking lot.

[5]Isopleth maps define spatial similarities, and enclose them with a boundary on the map. For example, weather maps often include isobars, a line connecting points having the same barometric pressure.

and potential area-unit fallacy faced by all area-unit data.

In addition, topographical representations of clusters have a number of disadvantages from the point of view of reader perception. The peaks and valleys make it difficult for the audience to see multiple hot spot areas, because a secondary peak may be hidden behind a larger primary peak. It is also difficult for a map reader to perceive the relationship between the topographical distribution and the position of other mapped objects (streets, buildings, parks). Topographical clusters of two separate variables (for example, abandoned buildings and drug offenses) are almost impossible to show clearly on the same map.

Pin map analysis

Pin map data, such as the locations of offenses, traffic accidents or known offender addresses, can provide a wealth of information, and statistical methods exist to organize that information into a useful form. Presentations at this workshop show how it is possible to analyze flow from one point to another (such as the distance from an offender's address to the site of a burglary), to identify high-activity "nodes" that attract a certain type of activity (such as a tavern or a dangerous intersection), and to rank-order locations according to frequency of activity (such as addresses with extremely high calls for service). None of this can be done with area-unit data.

However, frequency distributions of pin map or address data, by themselves, cannot define a particularly dense hot spot *area*. A single address with more crimes than any other address may, or may not, be located within the highest-density crime area on the map. Although it is an increasingly common practice (Joelson & Fishbine, 1980:256; Pierce, 1987; Sherman, et al., 1989; Uchida, 1990; Linnell, 1991) for crime analysts to define a single high-crime address as a "hot spot," and this designation may be useful for some applications, a hot spot address is not necessarily a Hot Spot Area.

In general, though areal and pin map analysis can provide answers to many other questions, neither can satisfy the basic requirement of Hot Spot Area analysis -- to identify high-density areas without regard to artificial boundaries. The predefined, arbitrary boundaries of areal analysis are an obstacle to the identification of such real high-density areas. On the other hand, the unit of analysis in a pin map is so detailed that "area" takes on a qualitatively different meaning, and density could reflect some unique characteristic of the particular location. Just as the use of arbitrary area boundaries as the unit of analysis may hide spatial patterns that cross those boundaries, the use of addresses as a unit of analysis may hide patterns of density that occur across groups of contiguous addresses.

Building Non-Arbitrary Areas from Pin Map Data

As Joelson and Fishbine (1980:250) suggested over a decade ago, we need to overcome the limitations of arbitrary area boundaries by utilizing "address-level aggregations." In other words, we need to find a technique for building non-arbitrary summary areas from pin data. There have been some attempts to do this.

Methods based on pin map information include the Nearest Neighbor significance test for the clustering of events within a boundary (Brantingham & Brantingham, 1984:222-223; Boots & Getis, 1988), Selden's (1976) use of the PATTER program to define a point that is the shortest distance from all other points as a "hot spot," and in general, centrographic analysis (Stephenson, 1980: 146-155; LeBeau, 1987; Ebdon, 1977), which can be used to find the imaginary point on the map that is closest to all events on the map (the mean center). These approaches begin with pin map data, but they also end there. None of them yields a summary, bounded, *area* that is calculated from individual points.

At least three research projects have explicitly confronted the problem of turning point data into area data. Joelson and Fishbine (1980:257-262) describe a project in which they drew isopleth maps based on assaults per square mile in concentric circles around a specific site in Minneapolis (Moby Dick's Bar). While this produced a bounded summary area from pin map information, it did not necessarily target the densest area on the map. Linnell (1991), Weisburd, Maher and Sherman (1991), and Maltz, Gordon and Friedman (1991:41) examine the proximity of address-level "hot spots" (address where a relatively high number of events occur), but do not develop an objective rule for combining them into hot spot areas.

Two of these studies, Weisburd, Maher and Sherman (1991), and Maltz, Gordon and Friedman (1991:41) rely on the judgement of experts to link together closely spaced hot spot address locations. In the Weisburd study, the expert was a member of the research team, who stood at the location of a hot address and recorded the physical characteristics of the area one block in every direction. In the Maltz study, people knowledgeable about the community were asked to draw freehand lines on a map around areas they considered to be high-crime or drug markets. Both methods defined hot spot areas according to the given expert's cognitive map, not according to the actual clustering of events.[6] Cognitive maps can provide valuable information, but they are not tools for summarizing information about spatial clustering.

In summary, the available spatial statistics for finding and defining the densest area on a map -- a Hot Spot Area -- do not measure up to the need for an automatic, objective tool that turns points into areas. In general, analysis methods based on relative crime density within arbitrary areal units suffer from aggregation biases, serious problems in interpretation, and are subject to the area-unit fallacy, while available analysis methods based on pin map data cannot identify particularly dense *areas*. The following section describes a spatial statistical tool that was developed in Illinois to answer the need to identify Hot Spot Areas.

[6] For a study comparing cognitive hot spots to Hot Spot Areas of officially-recorded crime, see the presentation by George Rengert elsewhere in this volume.

STAC HOT SPOT AREA ELLIPSES

When Illinois police jurisdictions first began to use computer mapping in the mid-1980s, they realized that they needed an objective and quick way to summarize mapped information. Several departments asked the Illinois Criminal Justice Information Authority for an automated, reliable method for identifying "hot spots" (the densest concentrations of incidents on a map).[7] In response to these requests, the Authority developed the STAC (Spatial and Temporal Analysis of Crime) package, a spatial statistical toolbox for law enforcement.[8] The initial development of STAC was carried out by Samuel Bates, under the direction of Associate Director Ed Maier, and was supported by a grant from the Bureau of Justice Statistics (BJS) of the U. S. Department of Justice (Garry, 1985).[9]

Although the current STAC package contains a number of statistical tools for spatial analysis, the Hot Spot Area was the original STAC capability, and is still the most used.[10] In one of the first public presentations of STAC Hot Spot Areas, to the board members of the Authority, Executive Director J. David Coldren showed examples of Hot Spot Area maps and commented (Coldren, 1986):

> One of the more promising techniques is what we've been calling our hot spot procedure. This is a procedure designed to look at all the crimes within a given area (such as a town) and search for the place of highest crime density, or where most of the crimes occurred. This is the so-called hot spot.

The initial STAC Hot Spot Area was not an ellipse, but a circle. the STAC Hot Spot program was an iterative search routine that identified one "hot circle," the densest circle on the map. Sam Bates's algorithm began by superimposing a grid over a study area, the boundaries of which were defined by the user. The grid could be rectangular or triangular. A circle, the radius of which was also defined by the user, was drawn around each node of the grid (figure 2). The nodes of the grid were spaced at one-half the radius, so that the circles overlapped, helping to assure that no cluster would be overlooked. The number of events in each circle was counted, and the circles were ranked by this number.

Then a new grid was created, the nodes of which were a distance of half the radius from the initial nodes, and the process of drawing overlapping circles and counting events within the circles was repeated for this grid (figure 3). The second list of rank-ordered circles was added to the first, and the circle having the greatest number of events was chosen as the hot spot (hot circle.) This initial hot spot routine was tested with data on several types of offenses, over a number of months, from five Illinois police departments (Bates, 1987).[11] The current

[7]The departments requesting hot spot capability were users of the Authority's PIMS (Police Information Management System), who were using the computer mapping software developed by PIMS and running on the HP 3000 mini-computer.

[8]Responding to such requests is one of the functions of the Authority's Statistical Analysis Center (SAC). SAC staff help people use data.

[9]Paul White, the Bureau of Justice Statistics grant monitor for the two STAC projects and for the Early Warning System project, had the vision to see the potential of STAC at its infancy. His enthusiasm about STAC, and his moral support, encouragement and advice were invaluable. Without Paul White, STAC would not exist.

[10]STAC currently has two modules, TIME and SPACE. SPACE currently includes Nearest Neighbor Analysis, Mean Center, radial searches around an address or other location, and another perspective on clustering, the Iso-crime. For more information about STAC, see the newsletter, *STACNews*.

[11]For testing, "flat files" were created by downloading data for a given month and offense type from the PIMS management system.

STAC Hot Spot Area algorithm has been enhanced, but it is still based on this basic foundation developed by Sam Bates.

Figure 2
Location of Lattice Points

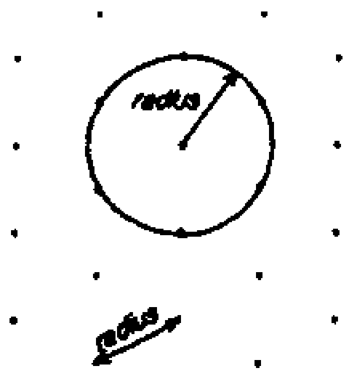
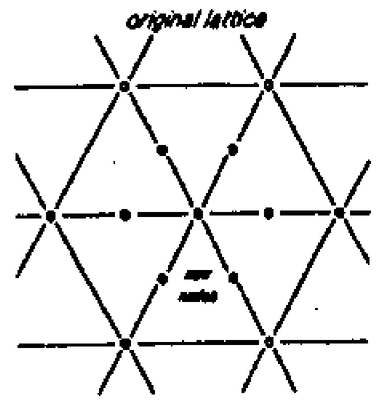


Figure 3
Location of New Nodes Relative to the Original Lattice



The potential of STAC seemed clear. However, there were two major drawbacks: it was a mainframe program, and it had never been tested in a police environment. Another grant from the Bureau of Justice Statistics (BJS) supported the translation of STAC from a mainframe to a PC program (written in C++), and testing in the Evanston, Illinois, Police Department.[12]

In addition, extensive testing of STAC by Sam Bates and the current author with data from the five test departments raised additional questions about the interpretation of the hot circle as the "densest area on the map."[13] Empirically, this seemed to be true -- STAC's hot circle always met the criterion of being denser (in incidents per square mile) than the rest of the map as a whole. However, every STAC hot spot search would find a hot circle, even in maps that were purposely constructed so that there was no real clustering. To demonstrate this, we created two kinds of maps with no clustering: maps in which the incidents were uniformly dispersed throughout the study area and maps in which the incidents were scattered randomly over the study area (using a program written by Sam Bates). In the random maps, the density of incidents in the hot spot circle found by STAC was always greater that the density outside of the hot circle (see Bates, 1987).

By definition, we knew that the higher density of hot circles in uniformly-dispersed or randomly-scattered maps represented only one permutation of random error. But this situation raised an important question: in general, when we do not know whether or not the incidents are really randomly distributed, how can we tell whether STAC's hot circle represents an area that is really more densely concentrated than other areas? Clearly, we needed a method to distinguish between "random" clustering and "real" clustering.

[12]Sharyn Barrington-Carlson was the programmer who re-wrote STAC in C++.

[13]For a full discussion of this issue, with examples from the test departments, see Bates (1987).

James Spring, a geographer who worked on the STAC project under the second BJS grant, realized that Nearest Neighbor Analysis (NNA) provided exactly such a significance test for spatial clustering, and added an NNA capability to STAC (see Spring, 1987).[14] Our advice to STAC users is to first test a dataset for significant clustering using NNA, and then, given clustering, use STAC to locate and define where the clusters are.[15] This is analogous to statistical significance versus association. STAC can describe where clustering occurs, assuming that it is occurring. However, NNA determines whether or not the clustering is significantly non-random. (For details of the "how hot is hot" question and the use of NNA in STAC, see Spring and Block, 1987, 1988a, 1988b, 1989.)

Several problems in STAC interpretation still remained. First, the hot circle would commonly overlap and share the same incidents with other circles, and ties in the rank order of circles were not unusual, making the choice of a single circle as "hottest" somewhat arbitrary. Second, while a circle was sometimes an accurate description of the real shape of a dense cluster of events, often the actual shape was irregular or elongated and not well described by a circle. To solve both problems, James Spring elaborated the original hot spot circle to produce the Hot Spot Ellipse. Beginning with the iterative search routine described above for hot circles, the ellipse algorithm first determines whether or not any of these hot circles overlap each other. If they do, all the incidents occurring in a group of overlapping hot circles become a "hot cluster" (figure 4). All the incidents occurring in any hot circle that stands by itself are also a hot cluster. These clusters are ranked by the number of incidents (an incident that occurs in two or more overlapping circles is counted only once).[16] Then STAC calculates the best-fitting standard deviational ellipse (Hsu & Tiedemann, 1968; Soot, 1975; Stephenson, 1980:151; LeBeau, 1987) for each cluster (figure 5).

Figure 4
Hot Clusters

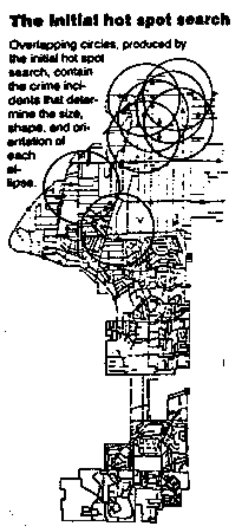


Figure 5
Hot Spot Ellipses

---

[14] James Spring's NNA program is written in FORTRAN. One of our priority STAC projects, given the resources, is to re-write it in C++ and integrate it with the rest of STAC. For more information on NNA, see Boots and Getis (1988).

[15] The NNA capability is currently limited in the STAC program to 400 cases. We hope to remedy this in the next version of STAC.

[16] The current version of STAC ranks Hot Spot Areas by the number of incidents in the cluster. In the enhanced version now being developed, the user will have a choice of alternative ranking criteria, including density (incidents per area of the ellipse).

At this point in STAC development, we felt that we needed feedback not only from users in the field, but also from experts in statistics and computer science. We were concerned about whether we could really say that the area encompassed by a STAC Hot Spot Ellipse was denser than any other possible ellipse, whether the STAC hot circle was really "hotter" than any other circle that could possibly be drawn on a given map. Helmut Epp, chair of the Computer Science Department at DePaul University and an expert in artificial intelligence, and two faculty members, volunteered to read through all the documentation on STAC, and to tell us if the methodology was appropriate to answer the questions at hand. After evaluating the STAC package, this team of experts concluded that the Hot Spot Ellipse is an appropriate method to locate and define the densest area(s) on the map. They pointed out that it is not a simple statistic but more like an artificial intelligence interactive procedure. STAC should be used to summarize and search for patterns, not as a parametric statistic. They also suggested that we continue to pursue the development of another tool for the STAC toolbox, the Isocrime (Poethig, 1989).

Compared to a circle, an ellipse can approximate much more closely the actual shape of a cluster of events. However, ellipses still cannot describe irregular clusters (L-shaped, for example). Another limitation of Hot Spot Ellipses is that they should not be compared across maps; each ellipse must be interpreted relative to a given hot spot area search, with a certain configuration of incidents within a certain border.[17] This is because the Hot Spot Ellipse algorithm considers all the incidents within the study area border when it identifies each hot spot area. High density is relative to the overall scatter of all these incidents. Therefore, the densest Hot Spot Area on a map in which most incidents are widely scattered may not be as dense as the densest Hot Spot Area on a map in which most incidents are clustered together.

To handle both problems (irregular clusters and map-to-map comparisons) we added another tool to the STAC toolbox -- the Isocrime (see Bates, 1987; Spring & Block, 1988a, 1988b).[18] The Isocrime describes a Hot Spot Area as a set of isopleth lines, each encompassing a given proportion of the total number of incidents on the map.[19] For example, an inner isopleth might enclose the densest 10 percent of incidents; a surrounding isopleth, the densest 20 percent; and an outer isopleth, the densest 30 percent. Isocrimes can take any shape, and can be compared from map to map. Two "Isocrime 10% lines" on two different maps mean exactly the same thing.

---

[17]Although STAC identifies the densest areas on a map regardless of arbitrary boundaries within the study area, the outer border around the study area is still a factor. A cluster of events that crosses this outer boundary may not be identified as a hot spot area as readily as a cluster within the study area. This border problem has been generally recognized in spatial point pattern analysis, but is so far unsolved (see Boots & Getis, 1988:49-45 for their discussion of "edge effects").

[18]The process of drawing Isocrimes is not yet fully automated in STAC. We expect to add it in the next version.

[19]To identify Isocrimes, we first find the most tightly clustered 30 percent of events within the study area (using Nearest Neighbor Analysis), and determine the mean center of these events (an imaginary point that is closer to all of these events than any other point). An enclosed irregular line is then drawn around the 10 percent that are closest to the mean center. Subsequent lines enclosing 20 and 30 percent may be drawn in other colors. Reader perception is improved if the inner (10%) isopleth is drawn in red, and the outer isopleths in cooler colors.

Hot Spot Isocrimes also have some disadvantages, compared to Hot Spot Ellipses. It is more difficult to compute their area than the area of an ellipse, which means that density (incidents within the hot spot boundary per area of the hot spot) is more difficult to obtain. Also, searches for multiple Isocrimes are not as straightforward as searches for multiple ellipses.[20] The Hot Spot Ellipse and the Isocrime, therefore, provide complementary perspectives to help the user locate and define Hot Spot Areas. Each is particularly appropriate in different situations, and answers different questions. To get the most accurate picture of a given map, the user should utilize the information provided by both.

STAC Hot Spot Ellipses were Beta tested by Wally Briefs in the Sunnyvale, California Department of Public Safety, Philip R. Canter in the Baltimore County Police Department, and Richard Block of Loyola University. These tests showed the potential of Hot Spot Area identification for crime analysis. For example, the Baltimore County Police Department produced Hot Spot Area maps of Halloween vandalism in the previous year, and passed them out to district officers, so that they could target prevention strategies in the current year (see Canter, 1993). However, STAC still was limited by input requirements that made it awkward to use, and by its dependence on plane geometry, which conflicted with the requirements of some mapping software for longitude/latitude data. Also, drawing STAC ellipses on a map was still awkward, because each was defined by only four points.

In addition, two difficulties in STAC interpretation began to be apparent. Extensive testing by James Spring of Hot Spot Ellipses of the same offense and the same area over time found cases in which the hot spot location suddenly moved to the other side of the map. These were not instances of displacement, but occurred because there was really more than one hot cluster on the map. A slight change in spatial configuration had caused the second-densest cluster to become the densest cluster. A more realistic procedure, we realized, would identify and plot all of the hot clusters, not just the densest. Also, field tests of user perception of Hot Spot Area maps found that people tend to think that larger ellipses indicate denser clusters of incidents than do smaller ellipses. (Actually, empirical testing indicates that the opposite is often the case.) We realized that to avoid misinterpretation, two enhancements would be necessary for STAC: the ability to find and plot multiple hot spots, and the calculation of Hot Spot Area density (per area and per population).

In 1991, therefore, Graham Taylor, the STAC analyst who succeeded James Spring, worked with programmer Terese Brand to make the following modifications and enhancements to STAC:

- Data entry requirements were made friendlier. STAC now could accept comma delimited data.
- Found a package that translates state-plane to longitude-latitude data (and vice versa), and incorporated it into STAC.[21] This meant that STAC could be used easily with the TIGER files and packages using them.
- The Hot Spot Ellipse module was revised so that more than one hot ellipse could be identified and mapped (the densest, second densest, and so on). The program now told the user how many Hot Spot Areas it has found, and the user then could choose how many of these to map.
- The capability of drawing a complete standard deviational ellipse for every Hot Spot Area, instead of plotting only four points of the ellipse, was added to STAC.

[20] A secondary Isocrime cluster can be identified by re-running the Isocrime routine on the remaining incidents after the first 30 percent have been eliminated from the dataset. To identify more than two Isocrimes on the same map, the 30 percent criterion should be changed to 20 percent or less. However, be sure to document the criterion you use, because the centroid of the tightest 10 percent often differs from the centroid of the tightest 20 or 30 percent.

[21] After a long but unproductive search by STAC staff, this package, a public-domain program published by the U. S. Bureau of Land Management (BLM), was finally found by STAC user Richard Block. Graham Taylor then wrote a "shell" program in C++, which took geocoded longitude/latitude data through the BLM translation to State/Plane, ran STAC, then returned 64 points defining each ellipse boundary to the BLM translator to produce longitude/latitude coordinates for mapping. Richard Block further modified this program, so that the dataset containing the points defining the ellipses is in the proper format for MapInfo boundary files. All of this is invisible to the user, who simply runs STAC and sees the results appear as ellipses on a map.

Density by area unit became possible after Graham Taylor wrote a program to calculate the area of a Hot Spot Ellipse. It also can be calculated by using mapping software such as MapInfo to determine the area within a Hot Spot boundary. A technique to calculate density per population or other Census block data was developed by Richard Block (see Block and Block, 1993).

In 1992, STAC won a finalist award from the Ford Foundation/ J. F. Kennedy School of Government, Harvard University, *Innovations in State and Local Government* program. Part of that money was used to hold this workshop, and another part was used to publish *STACNews*, but about half of the grant award was used to enhance STAC. Under the *Innovations* grant, we were able to hire Yunus Mohammed, a student in computer science at the University of Illinois at Chicago, to improve the accessibility of STAC to the average user. We also were able to upgrade our computer mapping software to support Mr. Mohammed in his efforts. The following are the changes made to STAC under the Innovations grant:

- Ø Menus have been added on top of STAC. There are two menus, a main menu and a spatial analysis (SPACE) sub-menu.
- Ø Coordinate conversions, which had been external to STAC, have now been incorporated into the STAC product. Before the enhancement, STAC required a user to run a separate "batch" program to prepare the data to run STAC. Now, this conversion process is "transparent" to the user. It is automatically done, without any thought or action on the user's part.
- Ø A multiple-run facility has been added to STAC. Users now can bring in more than one input file at a time, without having to start STAC again from the beginning. However, all input files must be in the format: Identifier, x-coord, y-coord.
- Ø Every STAC run now generates an audit trail -- a report of the files analyzed, statistical parameters requested, and results (including number of incidents, number of Hot Spot Areas identified, location and density of each area). With these reports, it will be easy to document and reproduce Hot Spot Area maps.

There are many things that still could be done to make STAC more useful, more accurate, and more accessible to users. As resources permit, we plan to make the following improvements to STAC:

- Ø The STAC program should be modularized (in other words, each sub-routine in the program should be individually documented and related to each other sub-routine). This is necessary so that the Authority staff can support STAC and narrow down the search for bugs. In addition, we need more detailed documentation of the C++ code in STAC.
- Ø We must correct bugs in Yunus Mohammed's Beta version of STAC. For example,, the Beta version does multiple hot spot searches, but writes over files when it does so.
- Ø All STAC programs and functions should be accessible through menus. Now, only some of them are.
- Ø More information is needed for the audit trail.
- Ø The user should be able to specify the location of fields. File formats should allow for the reverse of x and y coordinates, to agree with geographic practice.
- Ø The current Beta version is user-friendly for a specific application, but not for others. The new version should include input and output options for users of all applications.
- Ø The maximum number of cases for the hot spot search is 15,000; the maximum number of cases for NNA is 400. It should be possible to override these defaults.
- Ø NNA, which is now written in FORTRAN, should be integrated into the C++ program.
- Ø The algorithm for Isocrimes should be standardized and automated more completely.
- Ø More ellipse print options are needed. The user should be able to specify thresholds or priorities for printing multiple Hot Spot Area results. The user should have the option, for example, to print ellipses according to density (incidents per area in each Hot Spot Ellipse) or according to the number of incidents in the hot cluster. Also, the user should be able to specify a minimum number of incidents in a Hot Spot Area or a hot cluster for printed results.
- Ø Finally, we must address the border problem. That is, we should introduce a weight in the iterative search routine, that gives more weight to circles in which part of the circle is across the outer border (and about

which nothing is known). Boots and Getis (1988) describe such an algorithm.

These enhancements and improvements, completed and planned, will make STAC easier to use and easier to interpret, and therefore more accessible to officers, crime analysts, and community problem-solvers. However, since STAC is more of an interactive artificial intelligence tool than a cut-and-dried statistic, its use and interpretation is more an art than a science. The quality of STAC analysis, particularly analysis in real-life tactical and problem-solving situations, therefore, depends to some extent on the quantity of analysis. That is, the more STAC (or any other spatial statistical tool) is used to solve a practical law enforcement problem, the more it will be used, and the more effectively it will be used. Users will support each other with hints, war stories, and practical advice.

Most of the early users of STAC concentrated on the analysis of property crime, such as burglary, automobile theft or vandalism. To demonstrate the possibility of using STAC as an "information foundation for violence reduction," the Authority and the Chicago Police Department applied for and received a grant from the Bureau of Justice Statistics to develop such an information foundation for the identification of street gang violence crisis areas in Chicago. In the following section, we outline how STAC and a GeoArchive of law enforcement and community information have been combined to target an escalating high-risk situation, while there is still time to intervene and save lives.

AN EXAMPLE OF STAC AT WORK: THE EARLY WARNING SYSTEM PROJECT

The Early Warning System for Street Gang Violence, a joint project of the Chicago Police Department and the Illinois Criminal Justice Information Authority, is now underway in Chicago's Police Area Four, an area containing some of the riskiest neighborhoods for street gang violence in the city. The purpose of the project is to develop an automated early warning system for law enforcement, which will identify potential neighborhood crisis areas, areas that are at high risk for suffering a "spurt" of serious street gang-related violence and homicide. This early warning system will be based on a statistical model, which consolidates spatial information obtained from a variety of community and law enforcement sources and organized in a GeoArchive, and then uses automated Hot Spot Area identification and other geographic statistics as tools to target crisis neighborhoods.[22]

The Early Warning System project is founded on the premise that, since street gang violence is spatially anchored and occurs as the culmination of escalating incidents of revenge and retaliation, information compiled by community and neighborhood organizations, as well as by law enforcement, could be used to develop an "early warning system" of neighborhoods in crisis (see Spergel & Curry, 1990). Continuing escalation would then be prevented by crisis intervention and dispute mediation, using both internal community influences and external police support. Such a program has shown success in two pilot projects, in Chicago's Humboldt Park and in Philadelphia (Spergel, et al., 1984; Spergel, 1986), but requires the strong support of neighborhood agencies, churches, community groups, and the police department.

The Early Warning System project uses automated Hot Spot Area identification as a targeting tool. One of the most important defining characteristics of street gang violence is territoriality. However, the territorial unit most likely to become the site of serious violence is not a large or pre-defined area like a police district, but a smaller area defined by the gang (Suttles, 1972:187-201). In addition, violence changes over time, following

[22] A GeoArchive is a database containing address-based data from both law enforcement and community sources, linked to computer mapping capability, and set up so that it can be updated, maintained, mapped, analyzed and used by those who are developing and implementing strategies of crime reduction in the community. It is an "Information Foundation for Community Policing."

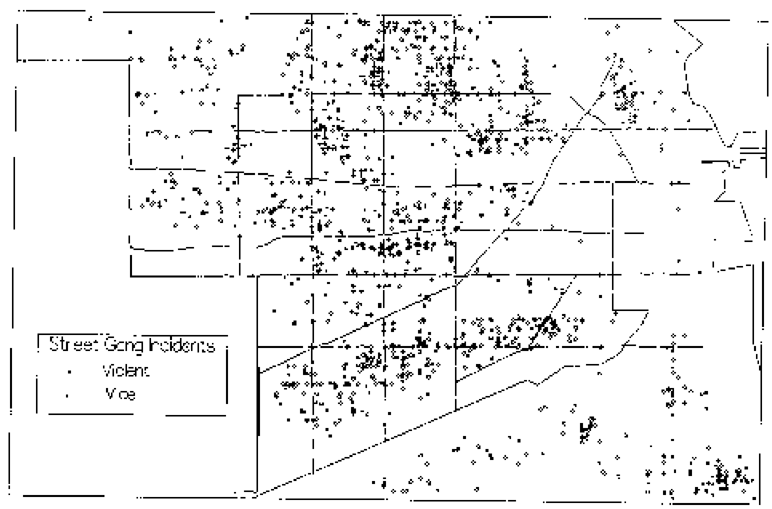
patterns of escalation, retaliation and revenge, often across a spatial border that may also change over time (see Weisburd, Maher & Sherman,1991:20). Therefore, a quick, objective statistical tool like the Hot Spot Area ellipses is vital, to target specific neighborhood areas that are at high risk of becoming a violence crisis area, so that crisis intervention, heading off the cycle of retaliation and retribution, is possible.

The datasets in the GeoArchive provide the basis for analysis of patterns in the study area. For example, the map in Figure 6 shows street gang-related criminal incidents in Area Four in 1992. Drug (vice) offenses are shown as "+" and violent offenses are shown as "o". The densest concentrations of street gang-related drug incidents are shown as light-line ellipses in Figure 7, and the densest concentrations of street gang-related violent incidents are shown as heavy-lines ellipses. This analysis shows that some neighborhoods have a Hot Spot Area of street gang-related drug offenses, others have a Hot Spot Area of street gang-related violent offenses, and still others have both. However, drug hot spots and "turf battle" hot spots are not necessarily in the same place (see Block & Block, 1993). If we are to develop a successful intervention strategy for street gang activity, we first must find out what kind of activity exists is a specific neighborhood. A city-wide program aimed at drug offending would have little impact in neighborhoods in which the preponderance of street gang-related incidents are not drug offenses, but violent offenses.

No matter how successfully we can develop an Early Warning System predictive model, however, it will not save lives unless there is a program that can use that information, and intervene in a potential crisis situation to save lives.

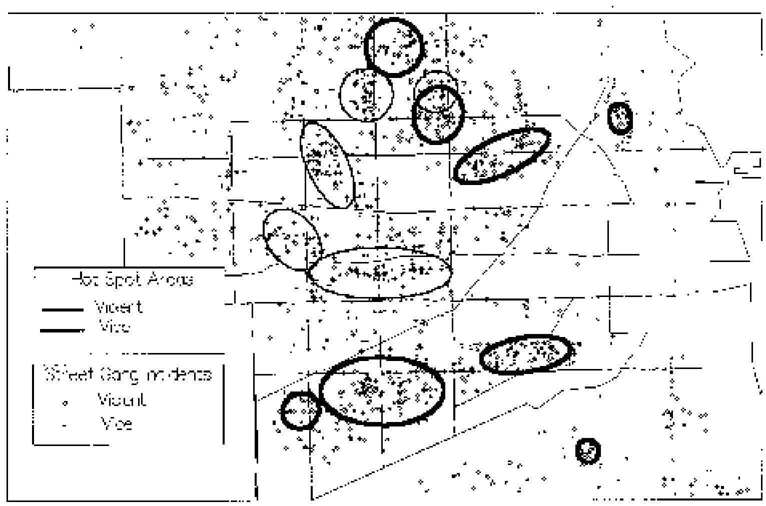
Fortunately, the Gang Violence Recduction Program has been established in District 10 of the Area Four study area, to do just that (CPD, 1993). The goal of the program is to reduce street gang-related violence by a collaborative multi-agency and community-law enforcement approach. The GeoArchive and STAC are being used as an information foundation for developing intervention strategies and responding to crisis situations.

Figure 6
Street Gang-Related Incidents, by Type
Vice Incidents and Violent Incidents



The long-range goal of the Early Warning System project is to develop a computer-assisted Early Warning System for neighborhoods in crisis, which will be transferrable to other police departments and communities throughout the country. Because of the importance of transferability of this project, the GeoArchive and computer-assisted Early Warning System are being be set up so that they can be updated, maintained, mapped, and analyzed by police analysts within the Area headquarters. We hope that the project will become a prototype for other neighborhoods and cities in developing similar automated systems for Hot Spot Area identification.

Figure 7
Hot Spot Areas



In summary, STAC is a database-driven, objective statistical tool that begins with individual pin map data and builds areas. STAC Hot Spot Areas reflect the actual scatter of events on a map, not arbitrary boundaries or pre-defined areas. Such statistical tools are necessary to control the "data overload" of law enforcement and community information, so that this information can become a foundation for community problem-oriented policing and other neighborhood-based problem-solving strategies. With a GeoArchive database of current, local-level law enforcement and community information, a theoretical framework and a statistical toolbox to organize this information, and cooperative community and law enforcement resources to intervene in a problem or crisis situation once identified, we can not only reduce levels of property crime, but also reduce the risk of injury and death from violence.

REFERENCES

Bates, Samuel (1987). *Spatial and Temporal Analysis of Crime*. Research Bulletin, April, 1987. Chicago: Illinois Criminal Justice Information Authority.

Block, Carolyn Rebecca (1989). The Application of Statistics to Criminal Justice Policy: A Populist Experiment. Invited paper, annual meeting of the American Association for the Advancement of Science, January 18, 1989.

___ (1990). Hot spots and isocrimes in law enforcement decision making. Paper presented at conference on Police and Community Responses to Drugs, University of Illinois at Chicago.

___ and Richard Block (1993). *Street Gang Crime in Chicago*. Research in Brief, NCJ 144782, December, 1993. Washington, D.C.: National Institute of Justice.

Block, Richard and Carolyn Rebecca Block (1993). Spatial analysis to prevent gang violence: The Area 4 project. Paper presented at the American Society of Criminology meetings, Phoenix.

Boots, Barry N. and Arthur Getis (1988). *Point Pattern Analysis*. Vol. 8, Scientific Geography Series (Grantlan Thrall, Ed.). Newbury Park: Sage Publications.

Brantingham, Paul J. and Patricia L. Brantingham (1984). *Patterns in Crime*. New York: Macmillan Publishing Co.

Canter, Philip (1993). Hot Spot Area Analysis in Baltimore County. *STACNews* vol.1,no.1 (Winter, 1993).

Choldin, Harvey M. and Dennis W. Roncek (1976). Density, population potential and pathology: A block-level analysis. *Journal of Public Data Use*, 4:19-29.

Coldren, J. David (1986). Application of Spatial and Temporal Analysis to Law Enforcement. Presentation to the Illinois Criminal Justice Information Authority, June 3, 1986.

CPD (Chicago Police Department) (1993). The gang violence reduction program. Pp. 3 & following. *CAPS News*. Chicago: Chicago Police Department. December, 1993.

Curtis, Lynn A. (1974). *Criminal Violence: National Patterns and Behavior*. (Chapter 7, Spatial Distributions.) Lexington, Mass: Lexington Books.

Ebdon, David (1977). *Statistics in Geography*. Great Britain: Basil Blackwell, Inc.

Felson, Marcus (1986). Predicting crime potential at any point on the city map. Pp. 127-136 in Robert M. Figlio, Simon Hakim and George F. Rengert (eds.), Metropolitan Crime Patterns. Monsey, New York: Criminal Justice Press, Willow Tree.

Garry, Steve (1985). Development of practical applications for the use of spatial and temporal data in crime analysis. Proposal submitted to the Bureau of Justice Statistics, January 29, 1985. Chicago: Illinois Criminal Justice Information Authority.

Georges-Abeyie, Daniel E. (1981). Studying black crime: A realistic approach. Pp. 97-109 in Brantingham

& Brantingham (eds.), Environmental Criminology. Beverly Hills:Sage.

Hsu, Shin-yi and Clifford E. Tiedemann (1968). A rational method of delimiting study areas for unevenly distributed point phenomena. *The Professional Geographer*, XX (6, November): 376-381.

Joelson, Mitchell R. and Glenn M. Fishbine (1980). The display of geographic information in crime analysis. Pp. 247-263 in Georges-Abeyie & Harries (eds.), *Crime: A Spatial Perspective*. New York: Columbia University Press.

King, Leslie J. (1969). *Statistical Analysis in Geography*. New Jersey: Prentice-Hall.

Kohfeld, Carol W. and John Sprague (1990). The organization of homicide events in time and space. Paper presented at annual meeting, American Society of Criminology.

LeBeau, James L. (1987). The methods and measures of centrography and the spatial dynamics of rape. *Journal of Quantitative Criminology*, 3(2,June):125-141.

Linnell, Deborah Diane (1991). *The Geographic Distribution of Hot Spots of Rape, Robbery, and Automobile Thefts in Minneapolis*. MA thesis, University of Maryland.

Lowe, John C. and S. Moryadas (1975). *The Geography of Movement*. Boston: Houghton Mifflin.

Maltz, Michael D., Andrew C. Gordon, Robert K. Lebailly, Warren Friedman, Marc Buslik and Matthias Casey (1987). Crime Statistics in the Small: Mapping Individual Events. Presented at the annual meeting of the American Society of Criminology.

___ , Andrew C. Gordon and Warren Friedman (1991). *Mapping Crime in its Community Setting: Event Geography Analysis*. New York: Springer-Verlag.

Pierce, Glenn L. (1987). The Development of a Police Problem/Crime Analysis Work Station. Presented at conference on law enforcement computer mapping, Illinois Criminal Justice Information Authority, September 9, 1987.

Poethig, Margaret (1988). Hot spots and Isocrimes. Pp. 11-13 in *the Compiler*, 9(2,fall). Chicago: Illinois Criminal Justice Information Authority.

Pyle, Gerald F. (1974). *The Spatial Dynamics of Crime*. Chicago: University of Chicago, Department of Geography. With Edward W. Hanten, Patricia Garstang Williams, Allen L. Pearson, II, J. Gary. Doyle, and Kwame Kwofie.

___ (1976). Spatial and temporal aspects of crime in Cleveland, Ohio. *American Behavioral Scientist*, 20(2):175-197.

Roncek, Dennis W. and Richard L. Block (1983). The effect of neighborhood change on homicide in Chicago. Paper presented at the American Society of Criminology meetings.

___ and Pamela A. Maier (1991). Bars, blocks, and crimes revisited: Linking the theory of routine activities to the empiricism of "hot spots." Manuscript.

Selden, Robert (1976). Recognizing the pattern of crime. *Energy and Technology Review*. Livermore, California: Lawrence Livermore Laboratory (Nov. 1976).

Sherman, Lawrence W., Patrick R. Gartin and Michael E. Buerger (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27(1):27-55.

Silk, John (1979). *Statistical Concepts in Geography*. Great Britain: George Allen & Unwin, LTD.

Smith, Christopher J. and Gene E. Patterson (1980). Cognitive mapping and the subjective geography of crime. Pp. 205-218 in Georges-Abeyie & Harries (eds.), *Crime: A Spatial Perspective*. New York: Columbia University Press.

Soot, Siim (1975). *Methods and Measures of Centrography: A Critical Study of Geographic Applications*. Occasional Publications of the Department of Geography, Luis E. Ortiz and Susan Gross (Eds.). Urbana, Illinois: Geography Graduate Student Association, University of Illinois.

Spelman, William (1988). *Beyond Bean Counting: New Approaches for Managing Crime Data*. Washington, D. C.: Police Executive Research Forum.

Spergel, Irving A. (1984). Violent gangs in Chicago: In search of social policy. *Social Service Review*, 58(2,June):199-226.

___ (1986). The violent gang problem in Chicago: A local community approach. *Social Service Review* (March): 94-130.

___ , Paul Colson, Francine Atwell and Sherri Dinguss (1984). Reduction of gang violence: A community development approach. Presented at the Second Annual Symposium on Street Work, University of Tubingen, West Germany, September.

___ and G. David Curry (1990). Strategies and perceived agency effectiveness in dealing with the youth gang problem. Pp. 288-309 in C. Ronald Huff (ed.), *Gangs in America*. Newbury Park: Sage.

Spring, James Vincent (1987). Spatial and temporal analysis of crime. Paper presented at the annual meeting of the Illinois Sociological Association.

___ and Carolyn Rebecca Block (1987). Spatial and temporal analysis of crime. Paper presented at annual meeting American Society of Geographers.

___ and Carolyn Rebecca Block (1988a). Finding crime hot spots: Experiments in the identification of high crime areas. Paper presented at the Midwest Sociological Society, March 25.

___ and Carolyn Rebecca Block (1988b). Hot spots and incident level data in law enforcement. Paper presented at the American Society of Criminology meetings, November 12.

___ and Carolyn Rebecca Block (1989). STAC Users Manual. Chicago: Illinois Criminal Justice Information Authority.

Stephenson, Larry K. (1980). Centrographic analysis of crime. Pp. 247-263 in Daniel E. Georges-Abeyie & Keith D. Harries (eds.), *Crime: A Spatial Perspective*. New York: Columbia University Press.

Suttles, Gerald D. (1972). *The Social Construction of Communities*. Chicago: University of Chicago Press.

Uchida, Craig (1990). NIJ Sponsors System to Speed Information to Police on Drug Hotspots. Pp. 8-9 in *NIJ Reports*, no. 221 (summer).

Weisburd, David, Lisa Maher and Lawrence Sherman (1991). Contrasting crime general and crime specific theory: The case of hot spots of crime. (With Michael Buerger, Ellen Cohn and Anthony Petrosino.) *Forthcoming in Advances in Criminological Theory*, Transaction